View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

ScienceDirect

Chinese Journal of Aeronautics 22(2009) 649-657

Chinese Journal of Aeronautics

www.elsevier.com/locate/cja

Accurate 3D Target Positioning in Close Range Photogrammetry with Implicit Image Correction

Zheng Jiandong[a,b], Zhang Liyan[a,*], Du Xiaoyu[a]

[a]*College of Mechanical and Electronic Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*

[b]*College of Electronic and Mechanical Engineering, Nanjing Forestry University, Nanjing 210037, China*

Received 25 October 2008; accepted 12 May 2009

Abstract

Accurate three-dimensional (3D) target positioning is of great importance in many industrial applications. Although various methods for reconstructing 3D information from a set of images have been available in the literature, few of them pay enough attention to the indispensable procedures, such as target extraction from images and image correction having strong influences upon the 3D positioning accuracy. This article puts forward a high-precision ellipse center (target point) extraction method and a new image correction approach which has been integrated into the 3D reconstruction pipeline with a concise implicit model to accurately compensates for the image distortion. The methods are applied to a copyright-reserved close range photogrammetric system. Real measuring experiments and industrial applications have evidenced the proposed methods, which can significantly improve the 3D positioning accuracy.

Keywords: close range photogrammetry; implicit camera model; image correction; feature extraction

1. Introduction

The ever-more-strong demand for high-precision three dimensional (3D) measurements has been prevailing in a variety of industrial applications, such as assembly of large equipment and registration of 3D coordinate data measured from different viewpoints etc. With the fast development of close range photogrammetry, reconstructing 3D information with one digital camera as an imaging sensor has caught close attention since using a hand-held camera to take images seems not only to be of greater convenience and economy than with coordinate measuring machines (CMMs) and range scanners, but also freed from their restrictions. In industrial applications, the positioning accuracy usually arouses the most serious concern. In general, the fundamental theory of recovering camera orientations of each shot as well as the 3D information from uncalibrated images has been extensively studied and well-understood[1-4]. However, the main interest lies in solving the geometric relationship of the multiple views. Relatively little attention has been paid to the indispensable procedures, such as target extraction from images, image correspondence establishment and image distortion correction, which are quite crucial to the 3D positioning accuracy. Many related publications in the field of multiple view geometry[5-6] assume that the image locations of the interested object, their correspondence across images, as well as the real camera model are known, but actually, these issues definitely deserve more detailed consideration, especially in industrial applications with requirements for relatively high accuracy. Although there are quite a number of articles focusing on image feature extraction[7-8] and image distortion correction[9-10], proper ways to incorporate them into the whole pipeline of the 3D reconstruction from images are still in dire need.

There are also some publications concerning performance estimation of 3D reconstruction from images. For instance, H. Mayer[11] discussed the robustness of automatic orientation, calibration, and disparity estimation in the reconstruction of image triplets. W. Forstner[12] analyzed the uncertainty of 3D reconstruction. E. Grossmann[13] studied the influence on precision of the number of images and the camera intrinsic parameters, etc. and compared the precision between calibrated and uncalibrated reconstructions. M. Sun, et

*Corresponding author. Tel.: +86-25-84891678.
E-mail address: zhangly@nuaa.edu.cn

Foundation items: National Natural Science Foundation of China (50875130); Doctoral Discipline Foundation of China (200802870016); Science Foundation of Jiangsu, China (BE2008136)

1000-9361 © 2009 Elsevier Ltd. Open access under CC BY-NC-ND license.
doi: 10.1016/S1000-9361(08)60154-5

al.[14] put forward a method to estimate precision of results of 3D reconstruction. However, these articles are deficient in shortage of an in-depth study of the accuracy of these algorithms. And what's more, nowhere in them can be found how to promote the camera precision and decrease the image point error.

This article investigates the way to accurately locate 3D target points from a small set of images taken with a hand-held camera. The target points are limited to visually salient coded or uncoded circular markers purposely disposed in the measuring field. After a brief introduction of the basic principles and the overall procedure of the solution in Section 2, an elaborate ellipse center (target point) extraction method is proposed in Section 3. Then, a new image correction procedure is put forward in Section 4, where a concise implicit model for image distortion compensation as well as the way to incorporate it into the whole 3D reconstruction pipeline is discussed in detail. The methods in Section 2 to Section 4 are incorporated into a copyright-reserved close range photogrammetry system. Section 5 presents real measuring experiments and industrial applications to show the proposed methods, which can improve the 3D positioning accuracy by approximately an order of magnitude if compared to the results in Ref.[15].

2. Basic Principles and Method Overview

Assume a set of images is captured with one camera in random poses. Two images in the set that cover a common area form an image pair. Let $\boldsymbol{X} = [X\ Y\ Z\ 1]^{\mathrm{T}}$ be the homogeneous coordinate of a target point in the 3D scene and $\boldsymbol{x} = [x\ y\ 1]^{\mathrm{T}}$ that of the corresponding image point. With the pinhole perspective camera model, the projection of $\boldsymbol{X}$ on the image plane is

$$\boldsymbol{x} = \boldsymbol{K}[\boldsymbol{R} \mid \boldsymbol{t}]\boldsymbol{X} = \boldsymbol{P}\boldsymbol{X} \tag{1}$$

where $\boldsymbol{R}$ is the rotation matrix, $\boldsymbol{t}$ the translation vector of the camera with respect to the world coordinate system, $\boldsymbol{P}$ a 3×4 perspective projection matrix, $\boldsymbol{K}$ a 3×3 matrix which contains camera intrinsic parameters

$$\boldsymbol{K} = \begin{bmatrix} f_x & s & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where f_x and f_y represent the focal length measured in pixels in the horizontal and vertical directions respectively, s is the skew factor, (u_0, v_0) the coordinates in pixel of the image center (the principal point).

For a 3D point $\boldsymbol{X}_i$, suppose $\boldsymbol{x}_{i1} = \boldsymbol{P}_1\boldsymbol{X}_i$ and $\boldsymbol{x}_{i2} = \boldsymbol{P}_2\boldsymbol{X}_i$ are the image locations of $\boldsymbol{X}_i$ in two images respectively. $\boldsymbol{x}_{i1}$ and $\boldsymbol{x}_{i2}$ are called a pair of corresponding image points with the relationship of $\boldsymbol{x}_{i1} \leftrightarrow \boldsymbol{x}_{i2}$. The epipolar geometry constraint is expressed as

$$\boldsymbol{x}_{i2}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{x}_{i1} = 0 \tag{3}$$

where $\boldsymbol{F}$ is the 3×3 fundamental matrix, which relates one image to the other geometrically.

The fundamental matrix $\boldsymbol{F}$ can be calculated based on the basic constraint in Eq.(3) by using the method recently proposed in Ref.[16] if the exact image locations and the correspondences of at least five points are given, i.e., $\boldsymbol{x}_{i1} \leftrightarrow \boldsymbol{x}_{i2},\ i = 1, 2, \cdots, n,\ n \geq 5$.

Since the initial value of the intrinsic parameter matrix $\boldsymbol{K}$ can be obtained from the manufacturer's camera specification, the essential matrix $\boldsymbol{E}$ can be calculated according to the following relationship:

$$\boldsymbol{E} = \boldsymbol{K}^{\mathrm{T}}\boldsymbol{F}\boldsymbol{K} \tag{4}$$

Align the world coordinate system to the first camera pose; then the projection matrix of the two images can be expressed as $\boldsymbol{P}_1 = \boldsymbol{K}[\boldsymbol{I} \mid \boldsymbol{0}]$ and $\boldsymbol{P}_2 = \boldsymbol{K}[\boldsymbol{R}_2 \mid \boldsymbol{t}_2]$ respectively. The rotation matrix $\boldsymbol{R}_2$ and the translation vector $\boldsymbol{t}_2$ of the second camera pose with respect to the first one can be determined by factoring matrix $\boldsymbol{E}$ using singular value decomposition (SVD) method. For details, refer to Ref.[16].

Having known both the intrinsic and the extrinsic parameters of the first two images, the 3D coordinates of all the interested points (including those not taking part in the above camera pose recovery process), which are seen in both images, can be calculated by using the stereo triangulation[17]. The bundle adjustment is then performed by the sparse Levenberg-Marquardt algorithm[18], which minimizes the re-rojection error with respect to all 3D points and camera parameters:

$$\min \sum_{ij} d(\boldsymbol{x}_{ij}, \boldsymbol{P}_j\boldsymbol{X}_i)^2 \tag{5}$$

where $\boldsymbol{P}_j\boldsymbol{X}_i$ is the predicted projection of the ith 3D point on the jth image; $\boldsymbol{x}_{ij}$ the detected location of $\boldsymbol{X}_i$ on the jth image and $d(\cdot)$ the Euclidean distance. Here j = 1, 2. Now the reconstruction of the first two images has been finished.

The associattion of camera pose and the 3D information contained in the jth image ($j \geq 3$) can be reconstructed incrementally till all the images have been treated [15,19].

3. Target Extraction from Image

From Section 2, it can be understood that to accurately reconstruct the 3D information and orient the images, extracting the exact image locations of the targets and robustly establishing the correspondences of them are firstly needed. They are the only input of the whole 3D reconstruction pipeline, on which the accuracy depends.

The targets used in the research context for camera pose recovery include the coded and uncoded markers[20] as shown in Fig.1. They are distributed in the measuring field. Precisely speaking, the center of the marker is the target point. Each coded marker has a

unique identity from the code band around the central solid white circle. According to the identity, the corresponding target points among images can be established directly[20]. An uncoded marker is a white solid circle with a concentric larger black dot as the background to make a high contrast making for detection. In our solution, the coded markers are mainly used to recover the camera poses following instructions in Section 2, so they can be regarded as auxiliary targets. Contrarily, the uncoded markers do not take part in the camera pose recovery process. The 3D coordinates of the uncoded targets are reconstructed by using stereo triangulation after the camera poses have been determined. The uncoded markers usually serve as the real targets in space positioning applications.

(a) Coded markers

(b) An uncoded marker

Fig.1 Two examples of coded markers and an uncoded marker.

As we know, the central white dot of both the markers inclusive of the coded and the uncoded is imaged to be an ellipse in a perspective projection. The representative subpixel methods for locating the ellipse center are the least-squares ellipse fitting[21] and the grey centroid method[22]. Based on the dual ellipse model, a simple operator for estimating the position and parameters of ellipse has been presented recently[23]. However, these methods are not directly adopted; instead, an improved subpixel algorithm is proposed here. First, color images are converted into grey ones and the Canny edge detector[24] is used to extract the edges from the images. Then, the Canny edges is corrected to achieve subpixel accuracy by using the polynomial surface fitting method. The surface equation can be written into

$$f(x,y)=k_1+k_2x+k_3y+k_4x^2+k_5xy+k_6y^2+k_7x^3+k_8x^2y+k_9xy^2+k_{10}y^3 \quad (6)$$

where (x, y) is the extracted edge's location using the Canny edge detector and $k_i(i=1,2,\cdots,10)$ the unknown coefficients.

For a given edge location (x, y) on a discrete digital image, $f(x, y)$ can be determined by the $N\times N$ neighboring elements of it. In practice, it is always to set $N=3$. The quick algorithm presented in Ref.[25] is adopted for obtaining the coefficients of the function $f(x, y)$.

According to Eq.(6), the gradient direction θ, the first order derivative and the second order derivative of θ are expressed by

$$\theta=\arctan(f'_y/f'_x)=\arctan\frac{k_3+k_5x+2k_6y+k_8x^2+2k_9xy+3k_{10}y^2}{k_2+2k_4x+k_5y+3k_7x^2+2k_8xy+k_9y^2} \quad (7)$$

$$\frac{\partial f}{\partial\theta}=(k_2+2k_4x+k_5y+3k_7x^2+2k_8xy+k_9y^2)\cos\theta+(k_3+k_5x+2k_6y+k_8x^2+2k_9xy+3k_{10}y^2)\sin\theta \quad (8)$$

$$\frac{\partial^2 f}{\partial\theta^2}=(6k_7\cos^2\theta+2k_8\sin(2\theta)+2k_9\sin^2\theta)x+(6k_{10}\sin^2\theta+2k_9\sin(2\theta)+2k_8\cos^2\theta)y+2(k_4\cos^2\theta+k_5\sin\theta\cos\theta+k_6\sin^2\theta) \quad (9)$$

The subpixel edge location (x_s, y_s) can be determined if it satisfies the following conditions:

$$\left.\begin{aligned}\left.\frac{\partial f(x,y)}{\partial\theta}\right|_{(x_s,y_s)}&\neq 0\\ \left.\frac{\partial^2 f(x,y)}{\partial\theta^2}\right|_{(x_s,y_s)}&=0\end{aligned}\right\} \quad (10)$$

Considering that subpixel edge detection can only be made on images with continuous intensity distribution, the intensity is obtained at the subpixel edge location (x_s, y_s) by the bilinear interpolation as follows

$$I(x_s,y_s)=(1-\alpha)(1-\beta)I(x,y)+\alpha(1-\beta)I(x+1,y)+(1-\alpha)\beta I(x,y+1)+\alpha\beta I(x+1,y+1) \quad (11)$$

where $\alpha=x_s-x$, $\beta=y_s-y$ and $I(x, y)$ is the intensity at (x, y).

The object region, i.e. the central white solid circle of each marker, is then divided into inner pixel region and edge pixel one. The intensities of the inner pixels are averaged to restrain the noise in the inner pixels. It is defined as

$$\bar{I}_{s_2}=\frac{1}{n_2}\sum_{i=1}^{n_2}I(x_{s_2,i},y_{s_2,i}) \quad (12)$$

where $I(x_{s_2,i}, y_{s_2,i})$ is the intensity at position $(x_{s_2,i}, y_{s_2,i})$ in the inner region and n_2 the number of the inner pixels.

The center of the coded marker is finally located by computing the following improved intensity weighted centroid

$$u' = \frac{\sum_{i=1}^{n_1} x'_{s_1,i} I(x'_{s_1,i}, y'_{s_1,i}) + \sum_{j=1}^{n_2} x_{s_2,j} \bar{I}_{s_2}}{\sum_{i=1}^{n_2} I(x'_{s_1,i}, y'_{s_1,i}) + n_2 \bar{I}_{s_2}} \tag{13}$$

$$v' = \frac{\sum_{i=1}^{n_1} y'_{s_1,i} I(x'_{s_1,i}, y'_{s_1,i}) + \sum_{j=1}^{n_2} y_{s_2,j} \bar{I}_{s_2}}{\sum_{i=1}^{n_1} I(x'_{s_1,i}, y'_{s_1,i}) + n_2 \bar{I}_{s_2}} \tag{14}$$

where $I(x'_{s_1,i}, y'_{s_1,i})$ is the intensity of the subpixel edge location $(x'_{s_1,i}, y'_{s_1,i})$ and n_1 the number of the edge pixels.

4. Integrated Image Correction

The linear pinhole camera model in Section 2 is only an approximation of the real camera projection. It is not valid when high-precision 3D reconstruction is required. Various sources of lens distortions, namely, radial, decentering and thin prism distortions[26] should be incorporated in a more sophisticated camera model. A representative one is[27]

$$\left.\begin{aligned} \tilde{x} &= x + x(k_1 r^2 + k_2 r^4) + \left[2p_1 xy + p_2(r^2 + 2x^2)\right] \\ \tilde{y} &= y + y(k_1 r^2 + k_2 r^4) + \left[p_1(r^2 + 2y^2) + 2p_2 xy\right] \end{aligned}\right\} \tag{15}$$

where $r = \sqrt{x^2 + y^2}$; (x, y) is the ideal (undistorted) normalized image coordinates following the linear model in Eq.(1); $(\tilde{x}, \tilde{y})$ the real (distorted) normalized image coordinates; k_1, k_2 are the radial and, p_1, p_2 the tangential distortion coefficients. The camera model in Eq.(15) corrects the ideal projection by the radial and tangential distortion components. Since the parameters k_1, k_2, p_1 and p_2 have their specific physical meanings, the model in Eq.(15) can be regarded as an explicit one. However, there is no analytic solution to the inverse mapping of Eq.(15) and a nonlinear search is required to recover (x, y) from $(\tilde{x}, \tilde{y})$. This is not easy to incorporate the model in the whole pipeline of 3D reconstruction from images. Enlightened by the work[28] and based on the extensive experiments, a more concise inverse mapping model can be put forward as follows:

$$\left.\begin{aligned} x &= \frac{1}{G}\left[\tilde{x} + \tilde{x}(c_1 \tilde{r}^2 + c_2 \tilde{r}^4)\right] \\ y &= \frac{1}{G}\left[\tilde{y} + \tilde{y}(c_1 \tilde{r}^2 + c_2 \tilde{r}^4)\right] \end{aligned}\right\} \tag{16}$$

where $G = (c_3 \tilde{r}^2 + c_4 \tilde{x} + c_5 \tilde{y} + c_6)\tilde{r}^2 + 1$; c_1, c_2, $\cdots$, c_6 are the distortion correction coefficients and $\tilde{r}^2 = \tilde{x}^2 + \tilde{y}^2$. Eq.(16) is an implicit model since the coefficients $c_1, c_2, \cdots, c_6$ are non-physical.

To integrate the procedure for solving $c_1, c_2, \cdots, c_6$ into the whole 3D reconstruction, the objective function in Eq.(5) is replaced by

$$\min \sum_{i=0}^{n} \sum_{j=0}^{m} \| \boldsymbol{x}_{ij}(c_1, c_2, \cdots, c_6) - \boldsymbol{x}_{ij}(f, d_x, d_y, u_0, v_0, \boldsymbol{R}_j, \boldsymbol{t}_j, \boldsymbol{X}_i) \| \tag{17}$$

in dealing with the first two images. Here $\boldsymbol{x}_{ij}(f, d_x, d_y, u_0, v_0, \boldsymbol{R}_j, \boldsymbol{t}_j, \boldsymbol{X}_i)$, which follows Eq.(1), is the predicted projection image coordinate of the ith 3D point $\boldsymbol{X}_i$ in the jth image; $\boldsymbol{x}_{ij}(c_1, c_2, \cdots, c_6)$, which follows Eq.(16), the corrected image coordinate of the distorted observed target points $\tilde{\boldsymbol{x}}_{ij}$.

The optimization problem concerning Eq.(17) is solved by the sparse Levenberg-Marquardt algorithm[18], which has proven to be rather too successful due to its adoption of the effective damping strategy that renders it able to converge promptly from a wide range of initial guesses. The objective function Eq.(17) is actually to minimize the squared distance $\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}$, where $\boldsymbol{\varepsilon} = \sum_{i=0}^{n} \sum_{j=0}^{m} (\boldsymbol{x}_{ij}(\tilde{\boldsymbol{P}}) - \boldsymbol{x}_{ij}(\tilde{\boldsymbol{P}}))$ and $\tilde{\boldsymbol{P}}$ is the parameter vector to be optimized. To reduce the overall number of parameters, the camera rotation is presented by a quaternion $\boldsymbol{q} = [q_1 \ q_2 \ q_3 \ q_4]$ of unit length. Therefore, the camera matrices and the correction coefficients of the jth image can be written in the following parameterized form:

$$\boldsymbol{a}_j = \left[f \ \ d_x \ \ d_y \ \ u_0 \ \ v_0 \ \ t_x^{(j)} \ \ t_y^{(j)} \ \ t_z^{(j)} \ \ q_1^{(j)} \ \ q_2^{(j)} \ \ q_3^{(j)} \ \ c_1 \ \ c_2 \ \ \cdots \ \ c_6\right]^{\mathrm{T}} \tag{18}$$

Also, the ith 3D point $\boldsymbol{X}_i$ can be parameterized by a vector like

$$\boldsymbol{b}_i = \left[X_x^{(i)} \ \ X_y^{(i)} \ \ X_z^{(i)}\right]^{\mathrm{T}} \tag{19}$$

Then the parameter vector $\tilde{\boldsymbol{P}} \in \mathbf{R}^M$ is shaped by concatenating all parameters:

$$\tilde{\boldsymbol{P}} = [\boldsymbol{a}_1^{\mathrm{T}} \ \ \boldsymbol{a}_2^{\mathrm{T}} \ \ \cdots \ \ \boldsymbol{a}_m^{\mathrm{T}} \ \ \boldsymbol{b}_1^{\mathrm{T}} \ \ \boldsymbol{b}_2^{\mathrm{T}} \ \ \cdots \ \ \boldsymbol{b}_n^{\mathrm{T}}]^{\mathrm{T}} \tag{20}$$

Given an estimated initial parameter $\tilde{\boldsymbol{P}}^0$, Levenberg-Marquardt algorithm searches for the best parameter vector $\tilde{\boldsymbol{P}}$, by iteratively solving a step vector $\boldsymbol{\delta}_{\mathrm{P}}$ with

$$(\mu \boldsymbol{I} + \boldsymbol{J}^{\mathrm{T}} \boldsymbol{J}) \boldsymbol{\delta}_{\mathrm{P}} = \boldsymbol{J}^{\mathrm{T}} \boldsymbol{\varepsilon} \tag{21}$$

where

$$\boldsymbol{J} = \frac{\partial (\sum_{i=0}^{n} \sum_{j=0}^{m} (\tilde{\boldsymbol{x}}_{ij}(\tilde{\boldsymbol{P}}) - \boldsymbol{x}_{ij}(\tilde{\boldsymbol{P}})))}{\partial \tilde{\boldsymbol{P}}}$$

is the Jacobian matrix and μ the damping term. After each iteration, the parameter vector $\tilde{\boldsymbol{P}}$ is updated by $\tilde{\boldsymbol{P}}+\boldsymbol{\delta}_{\mathrm{P}}$.

The entries in the Jacobian matrix can be calculated numerically by central differences. Notice that $\partial \boldsymbol{x}_{ij}/\partial \boldsymbol{a}_k=\boldsymbol{0}$, $\forall j\neq k$ and $\partial \boldsymbol{x}_{ij}/\partial \boldsymbol{b}_k=\boldsymbol{0}$, then $\forall j\neq k$ since the image point $\boldsymbol{x}_{ij}$ only depends on the parameters of the jth camera and ith 3D point $\boldsymbol{X}_i$. In addition, most partial differentiations of $\boldsymbol{x}_{ij}$ with respect to the entries in $\tilde{\boldsymbol{P}}$ is zero except for those with respect to c_1, c_2, $\cdots$, c_6. These facts mean that the Jacobian matrix is a highly sparse one, and the normal function Eq.(21) can be divided into smaller scale linear systems, which can be solved efficiently.

Having obtained the values of $c_1, c_2, \cdots, c_6$ from the first two images, all the distorted observed target points in all other images are corrected with the help of Eq.(16). After this, the other camera poses and the 3D information from the initially corrected jth image ($j\geq 3$) can be incrementally recovered. In the incremental procedure, the correction coefficients c_1, c_2, $\cdots$, c_6 should be repeatedly calculated according to the processed images and the subsequent images are corrected until the residual re-projection error of the bundle adjustment process drops under a limited threshold. When the incremental process is finished, the final bundle adjustment step optimizes all camera poses, all 3D target points and the intrinsic camera parameters including the distortion correction coefficients.

5. Experimental

The above-introduced methods have been incorporated into a copyright-reserved close range photogrammetry system named AutoLocator developed on the VC++ platform. The digital camera used in the experiments is a Nikon S1 with 4 256 × 2 848 pixel resolution.

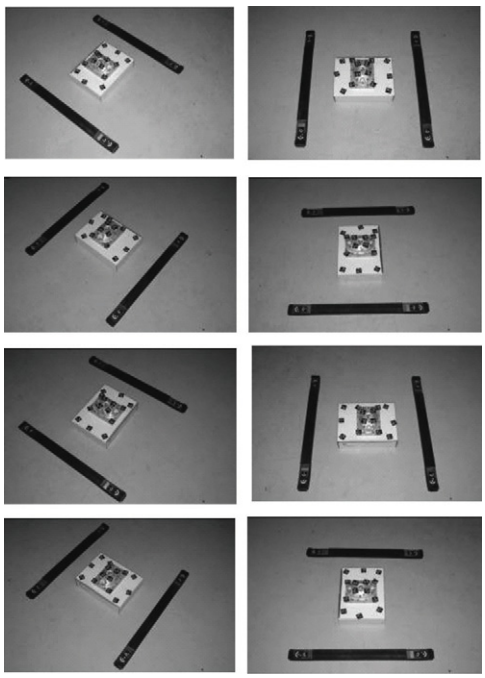
To analyze the accuracy of the recovered 3D point locations and restore the absolute 3D metric, two scale bars shown in Fig.2 are used as the length benchmarks. There are four coded points on each bar. The coded markers 0-1, 4-5, 2-3 and 6-7 constitute a marker pair respectively. As the distance between the centers of any one of marker pairs on the bars has been already given, the accuracy of 3D reconstruction can be evaluated by comparing the measured distance to the real distance of the marker pairs.

Fig.2 Scale bars with numbered coded points.

5.1. Influences of image target extraction algorithm

In this section, the performances of the proposed subpixel target extraction algorithm will be compared to the least-square ellipse fitting method[21], the grey centroid method[22]and the recently-presented dual ellipse operator[23]. Fig.3 shows some coded and uncoded markers that are disposed on a small die. A group of eight images around the die was taken by the hand-held camera Nikon S1 (Fig.3(a)) indoors. The 3D coded points and camera poses illustrated in Fig.3(b) are obtained with AutoLocator.

In this experiment, the image distortion is purposely left out, for the pinhole camera model is used from start to finish. Table 1 lists the evaluated accuracy results acquired by the different algorithms to extract the center point of the markers while keeping aspects the same.



(a) Image group

(b) 3D reconstructed information

Fig.3 3D reconstruction of targets on a die.

Table 1 Comparison of performance between four target extraction algorithms

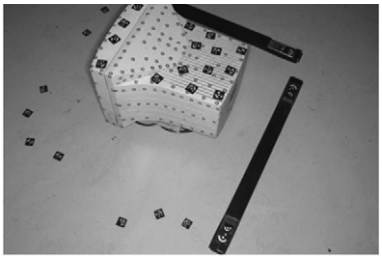
mm

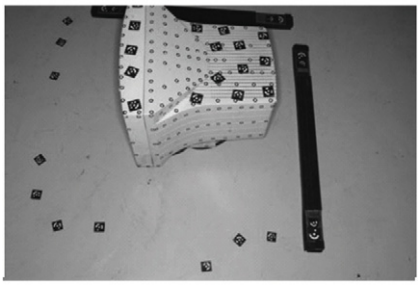
Marker pairs with known distances	Proposed algorithm by this article		Dual ellipse operator		Grey centroid method		Least-squares ellipse fitting	
	Estimated distance	Absolute error	Estimated distance	Absolute error	Estimated distance	Absolute error	Estimated distance	Absolute error
0-1 (611.898 0)	612.064 0	–0.166 0	612.057 4	–0.159 4	612.152 0	–0.254 0	612.156 0	–0.258 0
2-3 (611.800 0)	611.930 2	–0.130 2	612.025 5	–0.225 5	612.015 8	–0.215 8	612.015 1	–0.215 1
4-5 (545.940 0)	545.745 5	0.194 5	545.682 7	0.257 3	545.745 3	0.194 7	545.746 6	0.193 4
6-7 (545.823 0)	545.663 7	0.159 3	545.736 6	0.086 4	545.598 3	0.224 7	545.594 0	0.229 0
Average error	0.162 5		0.182 2		0.222 3		0.223 9	

Table 1 shows that the proposed target extraction algorithm obviously outperforms the other two traditional subpixel methods and so does the dual ellipse operator. The average length measuring error of the proposed algorithm is smaller than those of the traditional methods by approximately 27%, and the dual ellipse operator by about 11%. However, the value of the average error of 0.162 5 mm can in no way completely satisfy the high accuracy requirement called for under many industrial applications. Consequently, the following section will try to push ahead with the performances by way of the image distortion correction.

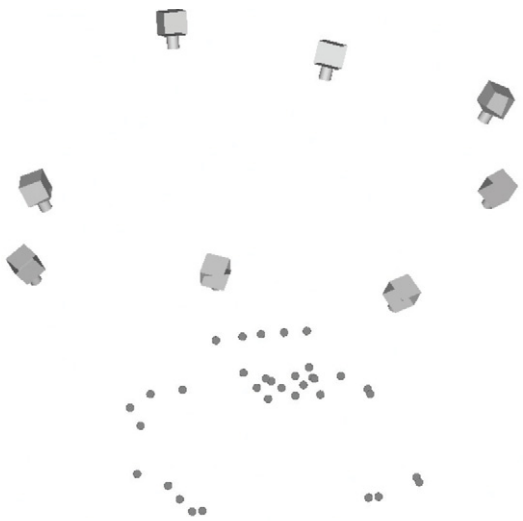
5.2. Effects of image correction algorithm

In this experiment, eight images are taken around a monitor with coded and uncoded targets as shown in Fig.4. Fig.4(a) shows two views in the image set. Fig.4(b) illustrates the recovered 3D coded points and camera poses in 3D space by using AutoLocator. Table 2 lists the image correction coefficients c_1, c_2, $\cdots$, c_6 found with Eq.(17). To demonstrate the effects of the proposed image correction algorithm, Table 3 illustrates the results from the comparison of accuracy between the cases with and without the image correction. The image target extraction algorithm used in both cases is the one presented in Section 3. It is clear that after introducing the implicit image correction into AutoLocator, the system accuracy has been up by an order of magnitude. This conclusion accords with the published results[15].





(a) An image pair



(b) Recovered 3D coded points and camera in 3D space

Fig.4 3D reconstruction of targets on a monitor.

Table 2 Implicit image correction coefficients

Coefficient	Value	Coefficient	Value
c_1	–0.016 072	c_4	0.000 000
c_2	0.000 075	c_5	–0.000 001
c_3	0.000 075	c_6	–0.016 105

Table 3 Performance comparison of with and without image correction

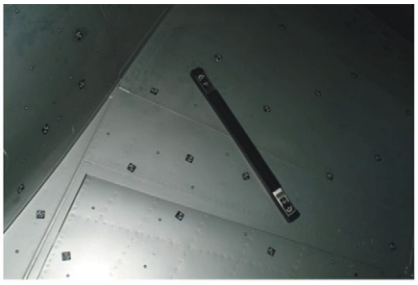
mm

Marker pair	Known distance	With image correction		Without image correction	
		Calculated distance	Absolute error	Calculated distance	Absolute error
0-1	611.898 0	611.899 3	–0.001 3	612.362 0	–0.464 0
2-3	611.800 0	611.759 7	0.040 3	611.605 5	0.194 5
4-5	545.940 0	545.960 1	–0.020 1	545.994 4	–0.054 4
6-7	545.823 0	545.837 6	–0.014 6	545.528 0	0.295 0
Average error		0.019 1		0.251 9	

5.3. A comprehensive industrial application

This section will present a real application accomplished by AutoLocator to demonstrate the overall accuracy level and the capability for dealing with a relatively large object. As shown in Fig.5, an airplane attached with quite a number of coded and uncoded markers is imaged by a hand-held camera from various positions and orientations with the aim to accurately position the uncoded points, which are to be used as data for aligning point clouds obtained by surface scanners (e.g. Ref.[29]) from different views. Fig.5(c) illustrates the 3D reconstruction results at the left wing and the joint region with fuselage. Up to 152 images are used for reconstructing the 3D targets and recovering all the camera poses within an approximate volume of 6 000 mm×1 500 mm×30 mm to be measured. The targets are attached to not only the upper side of the wing but also the lower side, so some images are taken from under the wing. Table 4 lists the given and the calculated distances between the marker pairs on the scale bars in this application.

(a) Scene when images were captured



(b) Three samples out of 152 images used for 3D reconstruction

(c) Recovered 3D target points (only the coded one) and 152 camera poses

Fig.5 A real target positioning application.

From Table 4, it is discovered that the average absolute error is 0.002 1 mm, and the average relative error 0.003 6 mm/m. This implies that with the incorporated methods suggested by this article, AutoLocator system achieves a quite high target positioning accuracy level; this is particularly true when the pro-

posed implicit image correction is simultaneously performed in the 3D reconstruction procedure, for it obviates the need for an accurate 3D calibration object. This would make it of great convenience in industrial applications.

Table 4 Accuracy evaluation in real application

Marker pair	Given distance/ mm	Calculated distance/mm	Absolute error/mm	Relative error/ ($mm \cdot m^{-1}$)
0-1	611.898 0	611.899 0	–0.001 0	0.001 6
2-3	611.800 0	611.803 4	–0.003 4	0.005 6
4-5	545.940 0	545.936 4	0.003 6	0.006 6
6-7	545.823 0	545.822 6	0.000 4	0.000 7

6. Conclusions

(1) With essential improvements of the 3D positioning accuracy, the proposed image target extraction algorithm, as is evidenced by the experiments, has an clear advantage over other three representative subpixel methods.

(2) The image distortion correction is highly critical to the 3D positioning accuracy, for the proposed implicit image correction model and the integrated 3D reconstruction pipeline are able to heighten the accuracy level of 3D positioning by an order of magnitude if compared to the simple linear model.

(3) Verified by real measuring experiments and industrial practices, the proposed methods have achieved high 3D reconstruction accuracy from a set of images, reaching an order of micron within a scope of several meters. The close range photogrammetry system AutoLocator has the potentiality of being widely used in industries.

References

[1] Hartley R I. Euclidean reconstruction from uncalibrated views. Proceedings of Conference on Computer Vision and Pattern Recognition. 1994; 908-912.

[2] Han M, Kanade T. Multiple motion scene reconstruction with uncalibrated cameras. IEEE Transactions on Pattern Analysis and Machine Intelligence 2003; 25(7): 884-894.

[3] Zhang Y, Zhang Z, Zhang J. Automatic measurement of industrial sheet metal parts with CAD data and nonmetric image sequence. Computer Vision and Image Understanding 2006(102): 52-59.

[4] Prakoonwit S, Benjamin R. 3D surface point and wireframe reconstruction from multiview photographic images. Image and Vision Computing 2007; 25(9): 1509-1518.

[5] Liu B, Yu M, Maier D, et al. An efficient and accurate method for 3D-point reconstruction from multiple views. International Conference of Computer Vision 2005; 65(3): 175-188.

[6] Heyden A, Astrom K. Euclidean reconstruction from constant intrinsic parameters. Proceedings of the International Conference on Pattern Recognition. 1996; 339-343.

[7] Zhang S C, Liu Z Q. A robust, real-time ellipse detector. Pattern Recognition 2005; 38(2): 273-287.

[8] Zhu Q, Wu B, Wan N. A sub-pixel location method for interest points by means of the Harris interest strength. The Photogrammetric Record 2007; 22(120): 321-335.

[9] Devernay F, Faugeras O. Straight lines have to be straight: automatic calibration and removal of distortion from scenes of structured environments. Machine Vision and Applications 2001; 13(1): 14-24.

[10] Thirthala S, Pollefeys M. The radial trifocal tensor: a tool for calibrating the radial distortion of wide-angle cameras. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005; 321-328.

[11] Mayer H. Robust orientation, calibration, and disparity estimation of image triplets. DAGM-Symposium, 2003; 281-288.

[12] Forstner W. Uncertainty and projective geometry. Handbook of Geometric Computing. Berlin: Springer, 2005; 493-535.

[13] Grossmann E. Uncertainty analysis of 3D reconstruction from uncalibrated views. Image and Vision Computing 2000(18): 685-696.

[14] Sun M, He R X, Wang D J. Precision analysis to 3D reconstruction from image sequences. ISPRS Workshop on Updating Geo-spatial Databases with Imagery & The 5th ISPRS Workshop on DMGISs. 2007; 141-146.

[15] Zheng J D, Zhang L Y, Zhou L, et al. 3D target location with one single hand-held CCD camera. Acta Aeronautica et Astronautica Sinica 2007; 28(6): 1521-1526. [in Chinese]

[16] Nister D. An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence 2004; 26(6): 756-770.

[17] Trucco E, Verri A. Introductory techniques for 3-D computer vision. New Jersey: Prentice Hall, 1998.

[18] Lourakis M I A, Argyros A A. The design and implementation of a generic sparse bundle adjustment software package based on the Levenderg-Marquardt algorithm. FORTH-ICS/TR-340, 2004.

[19] Zhang Z, Shan Y. Incremental motion estimation through local bundle adjustment. MSR-TR-01-54, 2001.

[20] Forbes K, Voigt A, Bodika N. An inexpensive, automatic and accurate camera calibration method. Proceedings of the Thirteenth Annual Symposium of the Pattern Recognition Association of South Africa. 2002.

[21] Fitzgibbon A, Pilu M, Fisher R B. Direct least square fitting of ellipses. IEEE Transactions on Pattern Analysis and Machine Intelligence 1999; 21(5): 476-480.

[22] Shortis M R, Clarke T A, Short T. A comparison of some techniques for the subpixel location of discrete target images. Proceedings of the SPIE: Videometrics III. 1994; 2350: 239-250.

[23] Ouellet J N, Hebert P. A simple operator for very precise estimation of ellipses. Proceedings of the Fourth Canadian Conference on Computer and Robot Vision. 2007; 28-30.

[24] Canny J. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 1986; 8(6): 679-698.

[25] Haralick R M. Digital step edges from zero crossing of

second directional derivatives. IEEE Transactions on Pattern Analysis and Machine Intelligence 1984; 6(1): 58-68.
[26] Weng J, Cohen P, Herniou M. Camera calibration with distortion models and accuracy evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence 1992; 14(10); 965-980.
[27] Heikkila J, Silven O. Calibration procedure for short focal length off-the-shelf CCD cameras. Proceedings of the 13th International Conference on Pattern Recognition. 1996; 166-170.
[28] Wei G Q, Ma S D. A complete two-plane camera calibration method and experimental comparisons. IEEE Proceedings of the 4th International Conference on Computer Vision. 1993; 439-446.
[29] Zhang H, Zhang L Y, Wang H T, et al. Surface measurement based on instantaneous random illumination. Chinese Journal of Aeronautics 2009; 22(3): 316-324.

Biographies:

Zheng Jiandong Born in 1975, he received M.S. degree from Nanjing University of Aeronautics and Astronautics in 2005, and then became a Ph.D. candidate there. His main research interests lie in computer vision and reverse engineering.
E-mail: njzg2006@163.com

Zhang Liyan Born in 1967, she received Ph.D. degree from Nanjing University of Aeronautics and Astronautics in 2001, and then became a professor there. Her main research interests lie in geometric design and computing, image-based 3D inspection and reverse engineering in product design.
E-mail: zhangly@nuaa.edu.cn